



ENTERED

RECEIVED
OIPE
P#7

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/770,875

DATE: 08/08/2002
TIME: 09:44:23

Input Set : A:\Nih42531.app
Output Set: N:\CRF3\08082002\I770875.raw

3 <110> APPLICANT: CIEPLAK, WITOLD
 5 <120> TITLE OF INVENTION: PERTUSSIN TOXIN GENE: CLONING AND EXPRESSION OF
 6 PROTECTIVE ANTIGEN
 8 <130> FILE REFERENCE: 2026-4253US7
 10 <140> CURRENT APPLICATION NUMBER: 09/770,875
 11 <141> CURRENT FILING DATE: 2001-01-26
 13 <150> PRIOR APPLICATION NUMBER: 07/311,612
 14 <151> PRIOR FILING DATE: 1989-02-15
 16 <150> PRIOR APPLICATION NUMBER: 07/542,149
 17 <151> PRIOR FILING DATE: 1990-06-22
 19 <150> PRIOR APPLICATION NUMBER: 08/483,326
 20 <151> PRIOR FILING DATE: 1995-06-07
 22 <150> PRIOR APPLICATION NUMBER: 09/128,911
 23 <151> PRIOR FILING DATE: 1998-08-04
 25 <160> NUMBER OF SEQ ID NOS: 28
 27 <170> SOFTWARE: PatentIn Ver. 2.1
 29 <210> SEQ ID NO: 1
 30 <211> LENGTH: 184
 31 <212> TYPE: DNA
 32 <213> ORGANISM: *Bordetella pertussis*
 34 <400> SEQUENCE: 1
 35 cccggggacag ggcggcgccc ggcggtcgct ggtccgcgcc ctggcgtgg tcctgccatc 60
 36 cggcgatg acgcatctt ccccccgcctt ggccgacgtt ctttatgtgc tggtaagac 120
 37 caatatggtg gtcaccagcg tagccatgaa gccgtatgaa gtcaccccgaa cgcggatgct 180
 38 ggtc 184
 41 <210> SEQ ID NO: 2
 42 <211> LENGTH: 61
 43 <212> TYPE: PRT
 44 <213> ORGANISM: *Bordetella pertussis*
 46 <400> SEQUENCE: 2
 47 Pro Gly Gln Gly Gly Ala Arg Arg Ser Arg Val Arg Ala Leu Ala Trp
 48 1 5 10 15
 50 Leu Leu Ala Ser Gly Ala Met Thr His Leu Ser Pro Ala Leu Ala Asp
 51 20 25 30
 53 Val Pro Tyr Val Leu Val Lys Thr Asn Met Val Val Thr Ser Val Ala
 54 35 40 45
 56 Met Lys Pro Tyr Glu Val Thr Pro Thr Arg Met Leu Val
 57 50 55 60
 60 <210> SEQ ID NO: 3
 61 <211> LENGTH: 17
 62 <212> TYPE: DNA
 63 <213> ORGANISM: *Bordetella pertussis*
 65 <220> FEATURE:

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/770,875

DATE: 08/08/2002

TIME: 09:44:23

Input Set : A:\Nih42531.app

Output Set: N:\CRF3\08082002\I770875.raw

66 <223> OTHER INFORMATION: Purine (P) R=G or A; Y=T or C; N=A, C, G, or T
 68 <220> FEATURE:
 69 <221> NAME/KEY: modified_base
 70 <222> LOCATION: (6)
 72 <220> FEATURE:
 73 <221> NAME/KEY: modified_base
 74 <222> LOCATION: (9)
 76 <220> FEATURE:
 77 <221> NAME/KEY: modified_base
 78 <222> LOCATION: (12)
 80 <220> FEATURE:
 81 <221> NAME/KEY: modified_base
 82 <222> LOCATION: (15)
 84 <400> SEQUENCE: 3
 W--> 85 atgaarccnt aygargt 17
 88 <210> SEQ ID NO: 4
 89 <211> LENGTH: 30
 90 <212> TYPE: PRT
 91 <213> ORGANISM: Bordetella pertussis
 93 <220> FEATURE:
 94 <223> OTHER INFORMATION: Xaa = Any amino acid; the 8th Val and 4th Pro are
 95 questionable.
 97 <220> FEATURE:
 98 <221> NAME/KEY: VARIANT
 99 <222> LOCATION: (15)
 101 <400> SEQUENCE: 4
 W--> 102 Asp Val Pro Tyr Val Leu Val Lys Thr Asn Met Val Val Thr Xaa Val
 103 1 5 10 15
 105 Ala Met Lys Pro Tyr Glu Val Val Pro Pro Arg Met Leu Val
 106 20 25 30
 109 <210> SEQ ID NO: 5
 110 <211> LENGTH: 4210
 111 <212> TYPE: DNA
 112 <213> ORGANISM: Bordetella pertussis
 114 <220> FEATURE:
 115 <221> NAME/KEY: CDS
 116 <222> LOCATION: (507)..(1313)
 118 <220> FEATURE:
 119 <221> NAME/KEY: CDS
 120 <222> LOCATION: (1356)..(2033)
 122 <220> FEATURE:
 123 <221> NAME/KEY: CDS
 124 <222> LOCATION: (2093)..(2485)
 126 <220> FEATURE:
 127 <221> NAME/KEY: CDS
 128 <222> LOCATION: (2500)..(2859)
 130 <220> FEATURE:
 131 <221> NAME/KEY: CDS
 132 <222> LOCATION: (2945)..(3625)

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/770,875

DATE: 08/08/2002
TIME: 09:44:23

Input Set : A:\Nih42531.app
Output Set: N:\CRF3\08082002\I770875.raw

134 <400> SEQUENCE: 5
 135 gaattcgtcg cctcgccctg gttcgccgtc atggcccca agggAACCGA ccccaagata 60
 137 atcgtcctgc tcaaccgcca catcaacgag gcgctgcagt ccaaggcggt cgtcgaggcc 120
 139 ttgcgcggcc aaggcgccac gccggtcatc gccacgcccgg atcagacccg cggcttcatc 180
 141 gcagacgaga tccagcgctg ggccggcggtc gtgcgcgaaa cggcgccaa gctgaagttag 240
 143 cagcgccggcc ctccaaacgctg ccatccccgt ccggccggca ccatccccca tacgtgttgg 300
 145 caaccgcggcc cgcgcattgcg tgcagattcg tcgtacaaaa ccctcgattc ttccgtacat 360
 147 cccgctactg caatccaaaca cggcatgaac gtccttcgg cgcaaaagtgc cgcatggta 420
 149 cccgtcaccg tccggaccgt gtcgacccccc ctgccatggc gtgatcccta aaataggcac 480
 151 catcaaaaacg cagaggggaa gacggg atg cgt tgc act cgg gca att cgc caa 533
 152 Met Arg Cys Thr Arg Ala Ile Arg Gln
 153 1 5
 155 acc gca aga aca ggc tgg ctg acg tgg ctg gcg att ctt gcc gtc acg 581
 156 Thr Ala Arg Thr Gly Trp Leu Thr Trp Leu Ala Ile Leu Ala Val Thr
 157 10 15 20 25
 159 gcg ccc gtg act tcg ccg gca tgg gcc gac gat cct ccc gcc acc gta 629
 160 Ala Pro Val Thr Ser Pro Ala Trp Ala Asp Asp Pro Pro Ala Thr Val
 161 30 35 40
 163 tac cgc tat gac tcc cgc ccg ccg gag gac gtt ttc cag aac gga ttc 677
 164 Tyr Arg Tyr Asp Ser Arg Pro Pro Glu Asp Val Phe Gln Asn Gly Phe
 165 45 50 55
 167 acg gcg tgg gga aac aac gac aat gtg ctc gac cat ctg acc gga cgt 725
 168 Thr Ala Trp Gly Asn Asn Asp Asn Val Leu Asp His Leu Thr Gly Arg
 169 60 65 70
 171 tcc tgc cag gtc ggc agc agc aac agc gct ttc gtc tcc acc agc agc 773
 172 Ser Cys Gln Val Gly Ser Ser Asn Ser Ala Phe Val Ser Thr Ser Ser
 173 75 80 85
 175 agc cgg cgc tat acc gag gtc tat ctc gaa cat cgc atg cag gaa gcg 821
 176 Ser Arg Arg Tyr Thr Glu Val Tyr Leu Glu His Arg Met Gln Glu Ala
 177 90 95 100 105
 179 gtc gag gcc gaa cgc gcc ggc agg ggc acc ggc cac ttc atc ggc tac 869
 180 Val Glu Ala Glu Arg Ala Gly Arg Gly Thr Gly His Phe Ile Gly Tyr
 181 110 115 120
 183 atc tac gaa gtc cgc gcc gac aac aat ttc tac ggc gcc gcc agc tcg 917
 184 Ile Tyr Glu Val Arg Ala Asp Asn Asn Phe Tyr Gly Ala Ala Ser Ser
 185 125 130 135
 187 tac ttc gaa tac gtc gac act tat ggc gac aat gcc ggc cgt atc ctc 965
 188 Tyr Phe Glu Tyr Val Asp Thr Tyr Gly Asp Asn Ala Gly Arg Ile Leu
 189 140 145 150
 191 gcc ggc gcg ctg gcc acc tac cag agc gaa tat ctg gca cac cgg cgc 1013
 192 Ala Gly Ala Leu Ala Thr Tyr Gln Ser Glu Tyr Leu Ala His Arg Arg
 193 155 160 165
 195 att ccg ccc gaa aac atc cgc agg gta acg cgg gtc tat cac aac ggc 1061
 196 Ile Pro Pro Glu Asn Ile Arg Arg Val Thr Arg Val Tyr His Asn Gly
 197 170 175 180 185
 199 atc acc ggc gag acc acg acc acg gag tat tcc aac gct cgc tac gtc 1109
 200 Ile Thr Gly Glu Thr Thr Glu Tyr Ser Asn Ala Arg Tyr Val
 201 190 195 200
 203 agc cag cat act cgc gcc aat ccc aac ccc tac aca tcg cga agg tcc 1157

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/770,875

DATE: 08/08/2002
TIME: 09:44:23

Input Set : A:\Nih42531.app
Output Set: N:\CRF3\08082002\I770875.raw

204	Ser	Gln	His	Thr	Arg	Ala	Asn	Pro	Asn	Pro	Tyr	Thr	Ser	Arg	Arg	Ser	
205	205							210					215				
207	gta	gcf	tcg	atc	gtc	ggc	aca	ttg	gtg	cgc	atg	gcf	ccg	gtg	ata	ggc	1205
208	Val	Ala	Ser	Ile	Val	Gly	Thr	Leu	Val	Arg	Met	Ala	Pro	Val	Ile	Gly	
209	220							225					230				
211	gct	tgc	atg	gcf	cgg	cag	gcc	gaa	agc	tcc	gag	gcc	atg	gca	gcc	tgg	1253
212	Ala	Cys	Met	Ala	Arg	Gln	Ala	Glu	Ser	Ser	Glu	Ala	Met	Ala	Ala	Trp	
213	235							240					245				
215	tcc	gaa	cgc	gcc	ggc	gag	gcf	atg	gtt	ctc	gtg	tac	tac	gaa	agc	atc	1301
216	Ser	Glu	Arg	Ala	Gly	Glu	Ala	Met	Val	Leu	Val	Tyr	Tyr	Glu	Ser	Ile	
217	250							255					260				265
219	gcf	tat	tcg	ttc	tagacctggc	ccagccccgc	ccaaactccgg	taattcaaca	gc								1355
220	Ala	Tyr	Ser	Phe													
222	atg	ccg	atc	gac	cgc	aag	acg	ctc	tgc	cat	ctc	ctg	tcc	gtt	ctg	ccg	1403
223	Met	Pro	Ile	Asp	Arg	Lys	Thr	Leu	Cys	His	Leu	Leu	Ser	Val	Leu	Pro	
224	270							275					280				285
226	ttg	gcc	ctc	ctc	gga	tct	cac	gtg	gcf	cg	gcc	tcc	acg	cca	ggc	atc	1451
227	Leu	Ala	Leu	Leu	Gly	Ser	His	Val	Ala	Arg	Ala	Ser	Thr	Pro	Gly	Ile	
228	290							295					300				
230	gtc	att	ccg	ccg	cag	gaa	cag	att	acc	cag	cat	ggc	agc	ccc	tat	gga	1499
231	Val	Ile	Pro	Pro	Gln	Glu	Gln	Ile	Thr	Gln	His	Gly	Ser	Pro	Tyr	Gly	
232	305							310					315				
234	cgc	tgc	gcf	aac	aag	acc	cgt	gcc	ctg	acc	gtg	gcf	gaa	ttg	cgc	ggc	1547
235	Arg	Cys	Ala	Asn	Lys	Thr	Arg	Ala	Leu	Thr	Val	Ala	Glu	Leu	Arg	Gly	
236	320							325					330				
238	agc	ggc	gat	ctg	cag	gag	tac	ctg	cgt	cat	gtg	acg	cgc	ggc	tgg	tca	1595
239	Ser	Gly	Asp	Leu	Gln	Glu	Tyr	Leu	Arg	His	Val	Thr	Arg	Gly	Trp	Ser	
240	335							340					345				
242	ata	ttt	gcf	ctc	tc	tac	gat	ggc	acc	tat	ctc	ggc	gaa	tat	ggc	ggc	1643
243	Ile	Phe	Ala	Leu	Tyr	Asp	Gly	Thr	Tyr	Leu	Gly	Gly	Glu	Tyr	Gly	Gly	
244	350							355					360				365
246	gtg	atc	aag	gac	gga	aca	ccc	ggc	ggc	gca	tcc	gac	ctg	aaa	acg	acg	1691
247	Val	Ile	Lys	Asp	Gly	Thr	Pro	Gly	Gly	Ala	Phe	Asp	Leu	Lys	Thr	Thr	
248	370							375					380				
250	ttc	tgc	atc	atg	acc	acg	cgc	aat	acg	gtt	caa	ccc	gca	acg	gat	cac	1739
251	Phe	Cys	Ile	Met	Thr	Thr	Arg	Asn	Thr	Gly	Gln	Pro	Ala	Thr	Asp	His	
252	385							390					395				
254	tac	tac	agc	aac	gtc	acc	gcc	act	cgc	ctg	ctc	tcc	agc	acc	aac	agc	1787
255	Tyr	Tyr	Ser	Asn	Val	Thr	Ala	Thr	Arg	Leu	Leu	Ser	Ser	Thr	Asn	Ser	
256	400							405					410				
258	agg	cta	tgc	gcf	gtc	ttc	gtc	aga	agc	ggg	caa	ccg	gtc	att	ggc	gcc	1835
259	Arg	Leu	Cys	Ala	Val	Phe	Val	Arg	Ser	Gly	Gln	Pro	Val	Ile	Gly	Ala	
260	415							420					425				
262	tgc	acc	agc	ccg	tat	gac	ggc	aag	tac	tgg	agc	atg	tac	agc	cgg	ctg	1883
263	Cys	Thr	Ser	Pro	Tyr	Asp	Gly	Lys	Tyr	Trp	Ser	Met	Tyr	Ser	Arg	Leu	
264	430							435					440				445
266	cgg	aaa	atg	ctt	tac	ctg	atc	tac	gtg	gcc	ggc	atc	tcc	gta	cgc	gtc	1931
267	Arg	Lys	Met	Leu	Tyr	Leu	Ile	Tyr	Val	Ala	Gly	Ile	Ser	Val	Arg	Val	
268	450							455					460				

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/770,875

DATE: 08/08/2002
TIME: 09:44:23

Input Set : A:\Nih42531.app
Output Set: N:\CRF3\08082002\I770875.raw

270	cat gtc agc aag gaa gaa cag tat tac gac tat gag gac gca acg ttc	1979
271	His Val Ser Lys Glu Glu Gln Tyr Tyr Asp Tyr Glu Asp Ala Thr Phe	
272	465 470 475	
274	gag act tac gcc ctt acc ggc atc tcc atc tgc aat cct gga tca tcc	2027
275	Glu Thr Tyr Ala Leu Thr Gly Ile Ser Ile Cys Asn Pro Gly Ser Ser	
276	480 485 490	
278	tta tgc tgagacgctt ccccaactcga accaccgccc cgggacaggg cgccgcggg	2083
279	Leu Cys	
280	495	
282	cggtcgcgc atg cgc gcc ctg gcg tgg ttg ctg gca tcc ggc gcg atg acg	2134
283	Met Arg Ala Leu Ala Trp Leu Leu Ala Ser Gly Ala Met Thr	
284	500 505	
286	cat ctt tcc ccc gcc ctg gcc gac gtt cct tat gtg ctg gtg aag acc	2182
287	His Leu Ser Pro Ala Leu Ala Asp Val Pro Tyr Val Leu Val Lys Thr	
288	510 515 520 525	
290	aat atg gtg gtc acc agc gta gcc atg aag ccg tat gaa gtc acc ccg	2230
291	Asn Met Val Val Thr Ser Val Ala Met Lys Pro Tyr Glu Val Thr Pro	
292	530 535 540	
294	acg cgc atg ctg gtc tgc ggc atc gcc gcc aaa ctg ggc gcc gcg gcc	2278
295	Thr Arg Met Leu Val Cys Gly Ile Ala Ala Lys Leu Gly Ala Ala Ala	
296	545 550 555	
298	agc agc ccg gac gcg cac gtg ccg ttc tgc ttc ggc aag gat ctc aag	2326
299	Ser Ser Pro Asp Ala His Val Pro Phe Cys Phe Gly Lys Asp Leu Lys	
300	560 565 570	
302	cgt ccc ggc agc agt ccc atg gaa gtc atg ttg cgc gcc gtc ttc atg	2374
303	Arg Pro Gly Ser Ser Pro Met Glu Val Met Leu Arg Ala Val Phe Met	
304	575 580 585	
306	caa caa cgg ccg ctg cgc atg ttt ctg ggt ccc aag caa ctc act ttc	2422
307	Gln Gln Arg Pro Leu Arg Met Phe Leu Gly Pro Lys Gln Leu Thr Phe	
308	590 595 600 605	
310	gaa ggc aag ccc gcg ctc gaa ctg atc cgg atg gtc gaa tgc agc ggc	2470
311	Glu Gly Lys Pro Ala Leu Glu Leu Ile Arg Met Val Glu Cys Ser Gly	
312	610 615 620	
314	aag cag gat tgc ccc tgaaggcgaa cccc atg cat acc atc gca tcc atc	2520
315	Lys Gln Asp Cys Pro Met His Thr Ile Ala Ser Ile	
316	625 630	
318	ctg ttg tcc gtc ctc ggc ata tac agc ccg gct gac gtc gcc ggc ttg	2568
319	Leu Leu Ser Val Leu Gly Ile Tyr Ser Pro Ala Asp Val Ala Gly Leu	
320	635 640 645	
322	ccg acc cat ctg tac aag aac ttc act gtc cag gag ctg gcc ttg aaa	2616
323	Pro Thr His Leu Tyr Lys Asn Phe Thr Val Gln Glu Leu Ala Leu Lys	
324	650 655 660 665	
326	ctg aag ggc aag aat cag gag ttc tgc ctg acc gcc ttc atg tcg ggc	2664
327	Leu Lys Gly Lys Asn Gln Glu Phe Cys Leu Thr Ala Phe Met Ser Gly	
328	670 675 680	
330	aga agc ctg gtc cgg gcg tgc ctg tcc gac gcg gga cac gag cac gac	2712
331	Arg Ser Leu Val Arg Ala Cys Leu Ser Asp Ala Gly His Glu His Asp	
332	685 690 695	
334	acg tgg ttc gac acc atg ctt ggc ata tcc gcg tat gcg ctc	2760

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/770,875

DATE: 08/08/2002
TIME: 09:44:24

Input Set : A:\Nih42531.app
Output Set: N:\CRF3\08082002\I770875.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; N Pos. 9

Seq#:4; Xaa Pos. 15

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/770,875

DATE: 08/08/2002
TIME: 09:44:24

Input Set : A:\Nih42531.app
Output Set: N:\CRF3\08082002\I770875.raw

L:85 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:0
L:102 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:0